

**Maîtrise B.P.E.  
Contrôle Terminal de MAB  
Statistique – 1 heure  
2ème session**

Tous documents autorisés – Calculatrices autorisées

**Question 1**

---

Soient  $\mathbf{x}' = [1 \ 1 \ 0]$  et  $\mathbf{y}' = [2 \ 1 \ 4]$  deux vecteurs de  $\mathbb{R}^3$ . Donner les coordonnées de  $\mathbf{z}$ , projeté orthogonal de  $\mathbf{y}$  sur  $\mathbf{x}$ .

**Question 2**

---

Soient  $\mathbf{x} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$  et  $\mathbf{y} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$  deux vecteurs de  $\mathbb{R}^2$ .

- a) Calculer la distance entre  $\mathbf{x}$  et  $\mathbf{y}$ .
- b) Chercher un vecteur  $\mathbf{z}$  orthogonal à  $\mathbf{x}$ .

**Question 3**

---

*Information* : Les matrices de la forme  $\begin{pmatrix} 1 & a & a \\ a & 1 & a \\ a & a & 1 \end{pmatrix}$  admettent pour valeurs propres  $\lambda_1 = 2a + 1$

et  $\lambda_2 = \lambda_3 = 1 - a$ .

On dispose des classements de 11 individus pour trois matières du certificat d'études (nouveau programme) :

|                 |    |   |   |   |   |   |   |   |   |    |     |
|-----------------|----|---|---|---|---|---|---|---|---|----|-----|
| <i>Maths</i>    | [1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11] |
| <i>Musique</i>  | [6 | 1 | 4 | 5 | 3 | 2 | 9 | 7 | 8 | 10 | 11] |
| <i>Sanskrit</i> | [2 | 6 | 5 | 3 | 4 | 1 | 8 | 9 | 7 | 10 | 11] |

On réalise une ACP normée du tableau 11 individus - 3 variables. La matrice des corrélations

est :  $\begin{pmatrix} 1 & 0.75 & 0.75 \\ 0.75 & 1 & 0.75 \\ 0.75 & 0.75 & 1 \end{pmatrix}$

- a) Donner l'inertie totale du nuage de  $\mathbb{R}^3$ .
- b) Donner les valeurs propres.
- c) Calculer le vecteur propre normé associé à la plus grande valeur propre.
- d) Un auditeur libre a eu des notes qui lui auraient donné les rangs 8 en maths, 2 en musique et 1 en sanscrit. Ecrire et calculer la coordonnée de cet auditeur sur le premier axe principal.

#### Question 4

On connaît pour 15 sportifs, notés de A à O, les performances dans trois épreuves mesurées à 18 ans et à 30 ans :

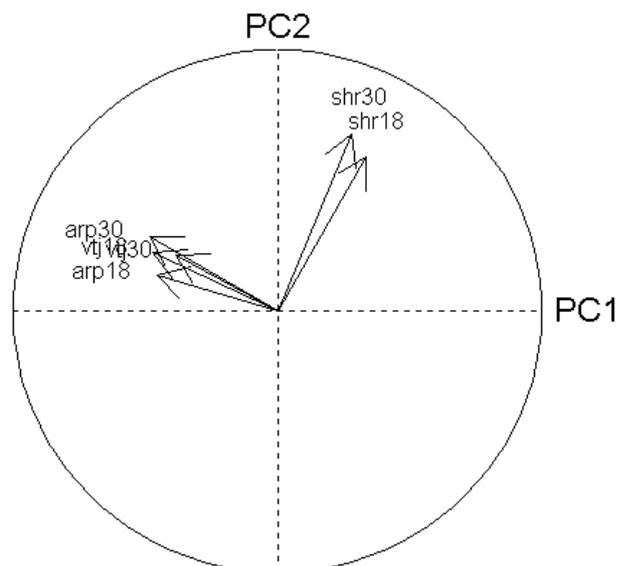
- vtj18 et vtj30 : détentés verticales, en cm (expression de la force explosive des membres inférieurs)
- arp18 et arp30 : forces du bras, en kg
- shr18 et shr30 : temps mis pour réaliser un test navette, en minutes.

Les données originales sont les suivantes :

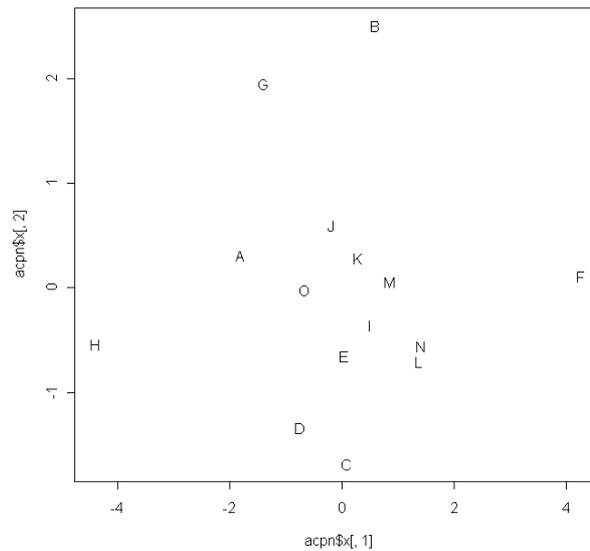
|   | VTJ18 | ARP18 | SHR18 | VTJ30 | ARP30 | SHR30 |
|---|-------|-------|-------|-------|-------|-------|
| A | 56    | 69.0  | 20.4  | 68    | 99.0  | 19.7  |
| B | 50    | 67.5  | 23.2  | 54    | 96.5  | 22.8  |
| C | 40    | 75.0  | 19.8  | 48    | 76.0  | 18.7  |
| D | 48    | 67.5  | 18.6  | 55    | 83.0  | 19.7  |
| E | 44    | 66.5  | 20.4  | 53    | 84.5  | 19.7  |
| F | 29    | 50.0  | 22.0  | 36    | 52.0  | 23.1  |
| G | 52    | 90.5  | 20.9  | 56    | 108.0 | 22.7  |
| H | 63    | 98.0  | 18.0  | 58    | 123.5 | 18.7  |
| I | 48    | 64.5  | 19.5  | 45    | 79.0  | 21.6  |
| J | 55    | 65.5  | 20.5  | 54    | 83.5  | 21.6  |
| K | 48    | 72.5  | 20.1  | 47    | 82.0  | 22.0  |
| L | 45    | 61.5  | 20.8  | 29    | 85.0  | 20.4  |
| M | 44    | 63.5  | 21.8  | 49    | 82.0  | 20.2  |
| N | 43    | 49.5  | 20.7  | 50    | 74.0  | 20.6  |
| O | 44    | 78.0  | 20.4  | 59    | 90.0  | 20.3  |

Une analyse en composantes principales normées a été réalisée sur les 15 sportifs et les 6 variables.

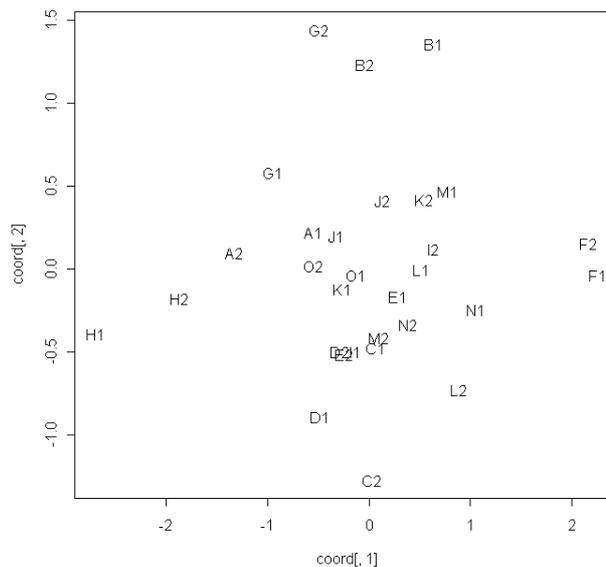
a) Interpréter le cercle des corrélations des axes 1 et 2.



b) On a représenté les individus, sur les deux premiers axes de l'analyse. Que pouvez vous dire des individus B,C, H et F ?



c) En réalité, un même individu a été mesuré deux fois : à l'âge de 18 ans (1) et à l'âge de 30 ans (2). En reprenant les quatre individus de la question précédente, réalisez une nouvelle analyse.



## Question 5

Pendant plusieurs années consécutives, on a observé le prix de vente moyen d'un certain type d'aliments pour le bétail<sup>1</sup>, exprimé en francs par kg.

| Année | 1972 | 1973 | 1974 | 1975 | 1976 | 1977 | 1978 |
|-------|------|------|------|------|------|------|------|
| Prix  | 5.87 | 6.38 | 6.44 | 6.28 | 6.68 | 7.02 | 7.08 |

Un modèle exprimant le prix en fonction de l'année est du style  $y = ax + b$ .

<sup>1</sup> P. Dagnelie "Théorie et méthodes statistiques. Exercices." Gembloux, 1994.

a) Quels sont les arguments permettant de rendre ce modèle probabiliste ?

b) Interpréter, biologiquement, les résultats ci-dessous :

$$y = 0.1839x - 356.7232$$

Analysis of Variance Table

Response: prix

|           | Df | Sum Sq  | Mean Sq | F value | Pr(>F)   |
|-----------|----|---------|---------|---------|----------|
| années    | 1  | 0.94723 | 0.94723 | 32.364  | 0.002340 |
| Residuals | 5  | 0.14634 | 0.02927 |         |          |

## Question 6

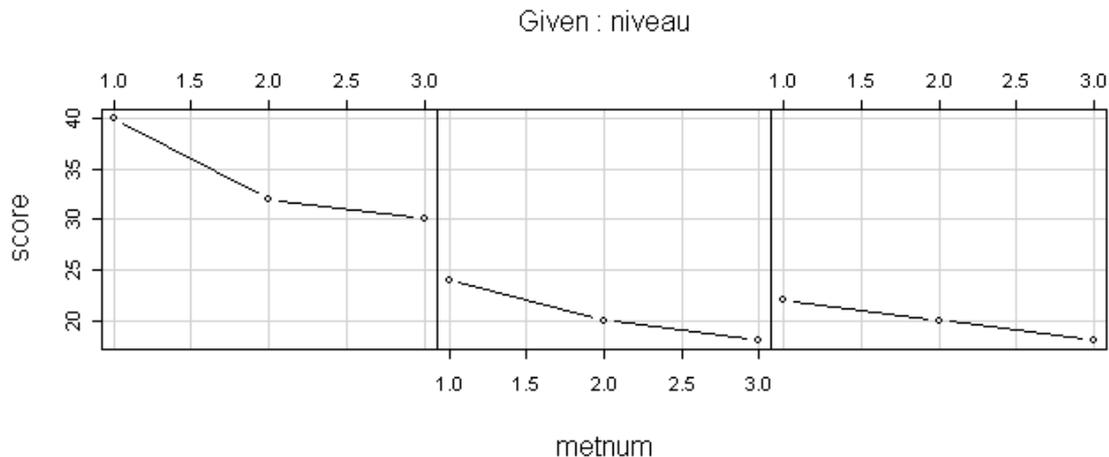
Un chercheur en sciences de l'éducation propose trois méthodes d'enseignement [method]. Il se demande si celles-ci sont également efficaces pour des élèves ayant des niveaux différents de lecture [niveau] : fort, moyen, faible.

Puis, il réalise un épreuve et donne un note [score] à chaque élève.

Les résultats sont consignés dans le tableau ci-dessous.

| Niveau de lecture | méthode 1 | méthode 2 | méthode 3 |
|-------------------|-----------|-----------|-----------|
| fort              | 40        | 32        | 30        |
| moyen             | 24        | 20        | 18        |
| faible            | 22        | 20        | 18        |

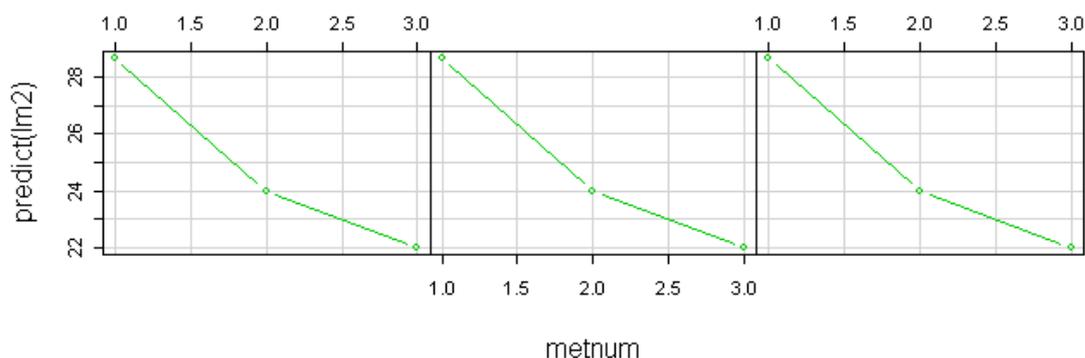
Représentation des notes obtenues par méthode, pour chacun des trois niveaux de lecture.



Trois modèles sont retenus pour analyser la situation.

**[A]** le modèle "effet méthode" :  $\hat{y} = c_1\mathbf{K}_1 + c_2\mathbf{K}_2 + c_3\mathbf{K}_3$

Given : niveau

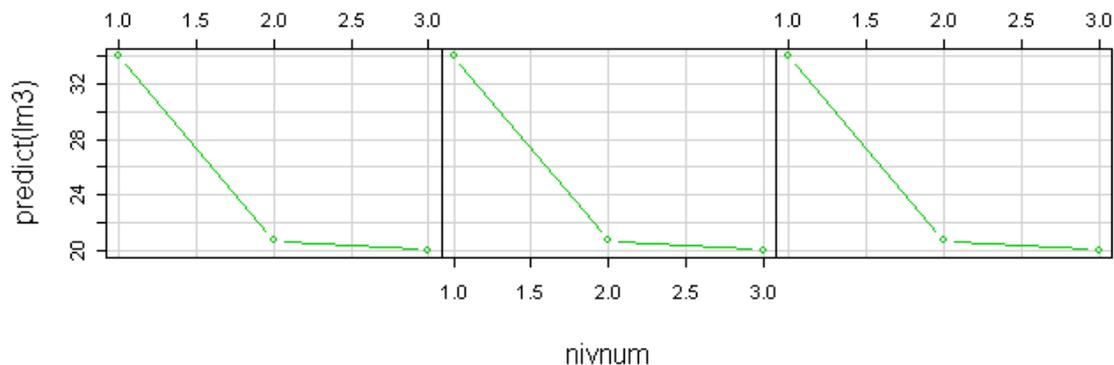


Response: score

|           | Df | Sum Sq | Mean Sq | F value | Pr(>F) |
|-----------|----|--------|---------|---------|--------|
| method    | 2  | 70.22  | 35.11   | 0.5448  | 0.6061 |
| Residuals | 6  | 386.67 | 64.44   |         |        |

**[B]** le modèle "effet niveau de lecture"  $\hat{y} = d_1L_1 + d_2L_2 + d_3L_3$

Given : method

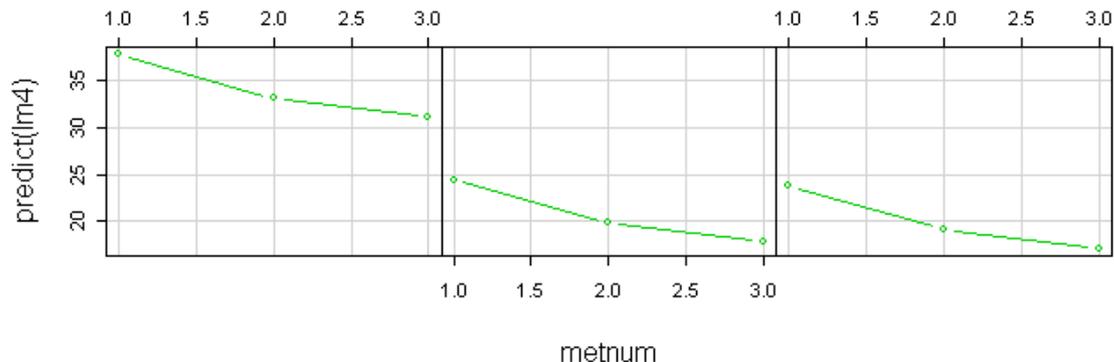


Response: score

|           | Df | Sum Sq | Mean Sq | F value | Pr(>F)      |
|-----------|----|--------|---------|---------|-------------|
| niveau    | 2  | 374.22 | 187.11  | 13.581  | 0.005923 ** |
| Residuals | 6  | 82.67  | 13.78   |         |             |

**[B]** le modèle "double effet"  $\hat{y} = c_1K_1 + c_2K_2 + c_3K_3 + d_1L_1 + d_2L_2 + d_3L_3$

Given : niveau



Response: score

|           | Df | Sum Sq | Mean Sq | F value | Pr(>F)      |
|-----------|----|--------|---------|---------|-------------|
| method    | 2  | 70.22  | 35.11   | 11.286  | 0.022662 *  |
| niveau    | 2  | 374.22 | 187.11  | 60.143  | 0.001036 ** |
| Residuals | 4  | 12.44  | 3.11    |         |             |

- 1) Comment obtient les valeurs  $c_1$ ,  $c_2$  et  $c_3$  du modèle **[A]** ?
- 2) Choisir, parmi les modèles **[A]**, **[B]** et **[C]**, celui qui vous paraît décrire au mieux les données.

§ § § § §